Proteins in Materials Science

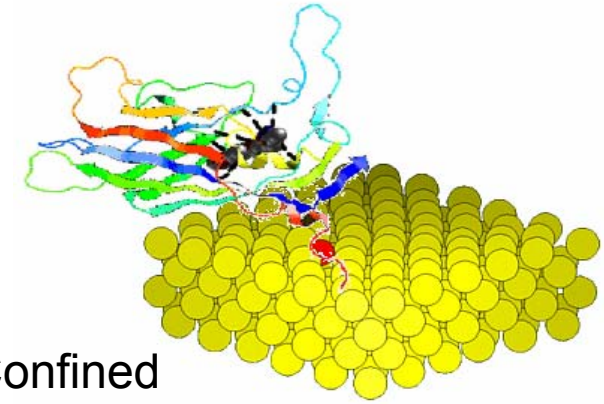
November 13th, 2004

MELIK DEMIREL, PhD
Assistant Professor and Pearce Development Professor
School of Engineering
Pennsylvania State University

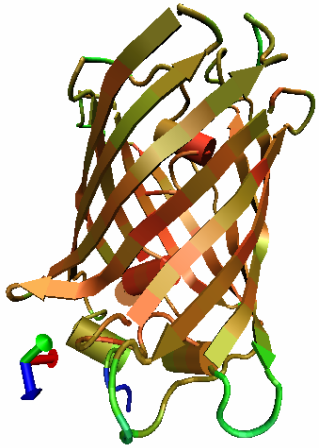
Protein Theory and Simulations

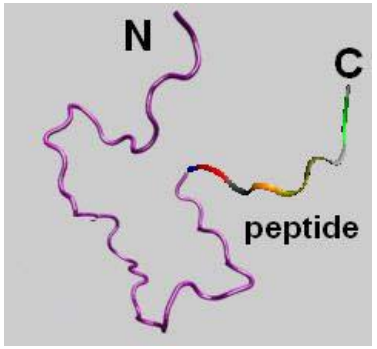
Protein modeling : Analytical approaches (GNM), chemical kinetic equations, molecular simulations (AMBER, XMD), ab initio calculations (VASP) (in collaboration with Materials Research Institute, Penn State)

LION-XM
128 node super computers (cost ½ million $)



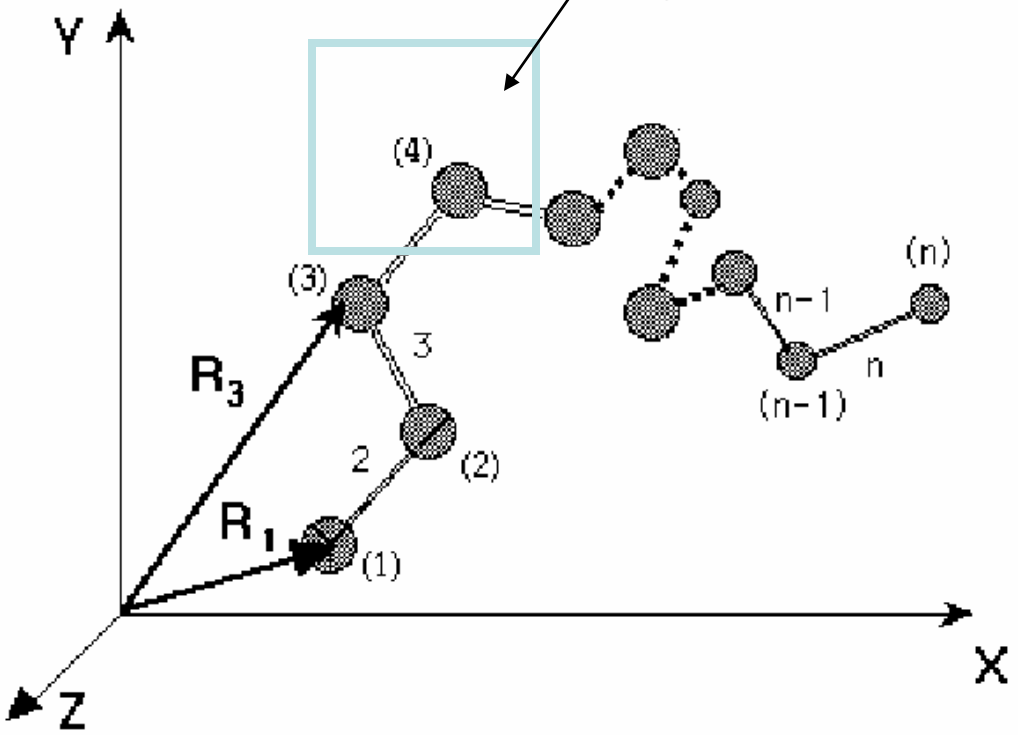
Confined Environments: Protein-Surface Interactions



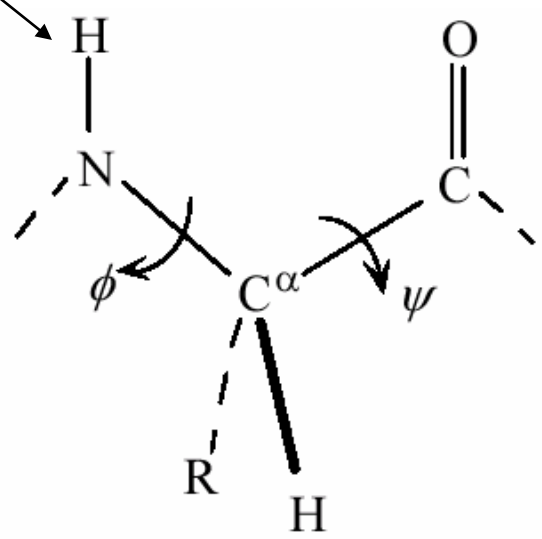


Protein Design

Repeating unit



Representation of a polymer chain



Proteins are heteropolymers made of amino acids

MOLECULAR DYNAMICS SIMULATIONS

- A deterministic method based on the solution of Newton's equation of motion

$$\mathbf{F_i = m_i a_i}$$

for the ith particle; the acceleration at each step is calculated from the negative gradient of the overall potential, using

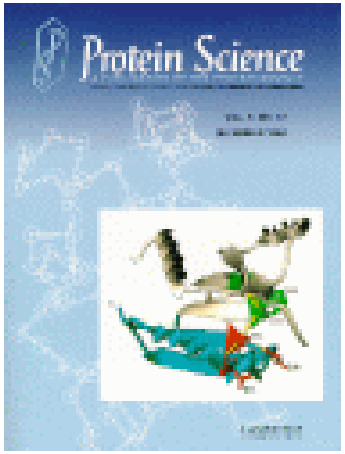
$$F_i = - \text{grad } V_i - = - \nabla \mathbf{V_i}$$

LIMITATIONS OF MD

- Full atomic representation → noise
- Empirical force fields → limited by the accuracy of the potentials
- Time steps constrained by the fastest motion (bond stretching of the order of femtoseconds)
- Inefficient sampling of the complete space of conformations
- High computational cost: Limited to small proteins (100s of residues) and short times (subnanoseconds)

ALTERNATIVE METHOD

COARSE GRAIN SIMULATIONS



GAUSSIAN NETWORK MODEL (GNM)

Demirel, M.C. (with others)
Protein Science,
December 1998

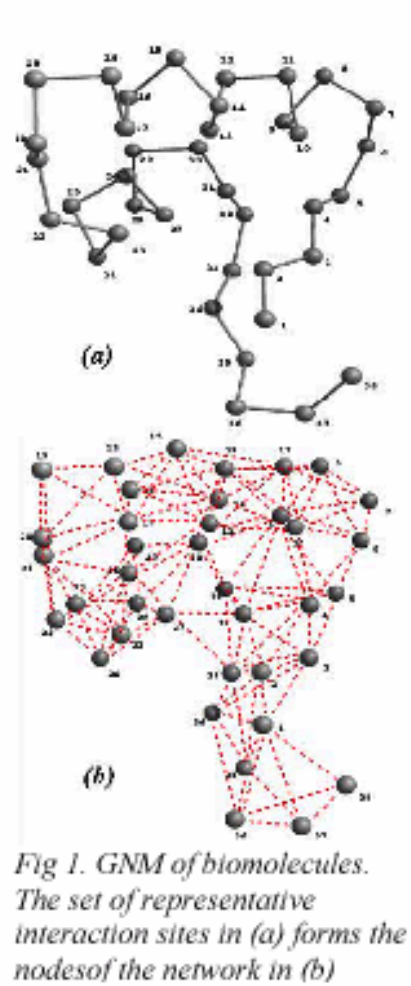


Fig 1. GNM of biomolecules. The set of representative interaction sites in (a) forms the nodesof the network in (b)

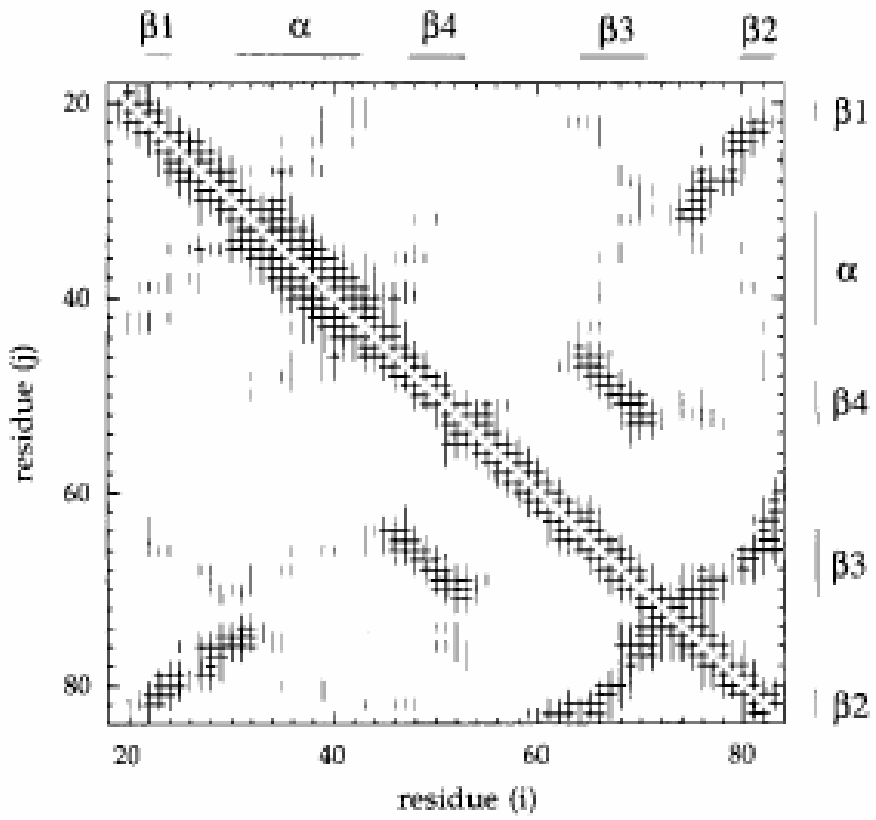
$$\mathcal{H} = \frac{1}{2}\gamma[\Delta\mathbf{R}^{\mathrm{T}}(\Gamma \otimes \mathbf{E})\Delta\mathbf{R}],$$

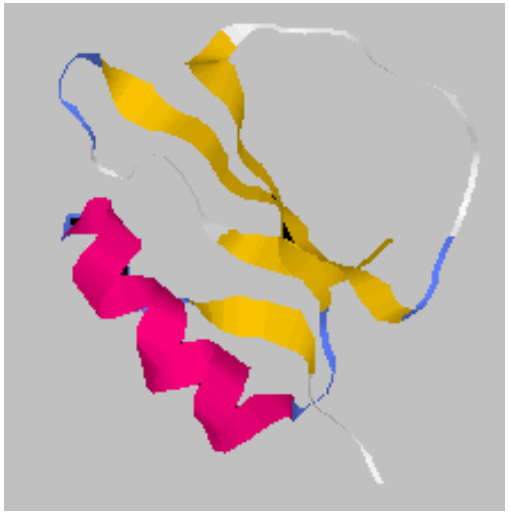
$$\Gamma^{-1}=(\gamma/3k_BT)*<\Delta R\Delta R^T>$$

$$A = -k_BT \ln Z_N = -(3k_BT/2)\ln[(\pi/\gamma^*)^{N-1}\det(\Gamma^{-1})],$$

Γ: Connectivity matrix, ΔR: fluctuation of each residue

CONTACT MAP (CONNECTIVITY MATRIX)





Chymotrypsin inhibitor-2 64 residues

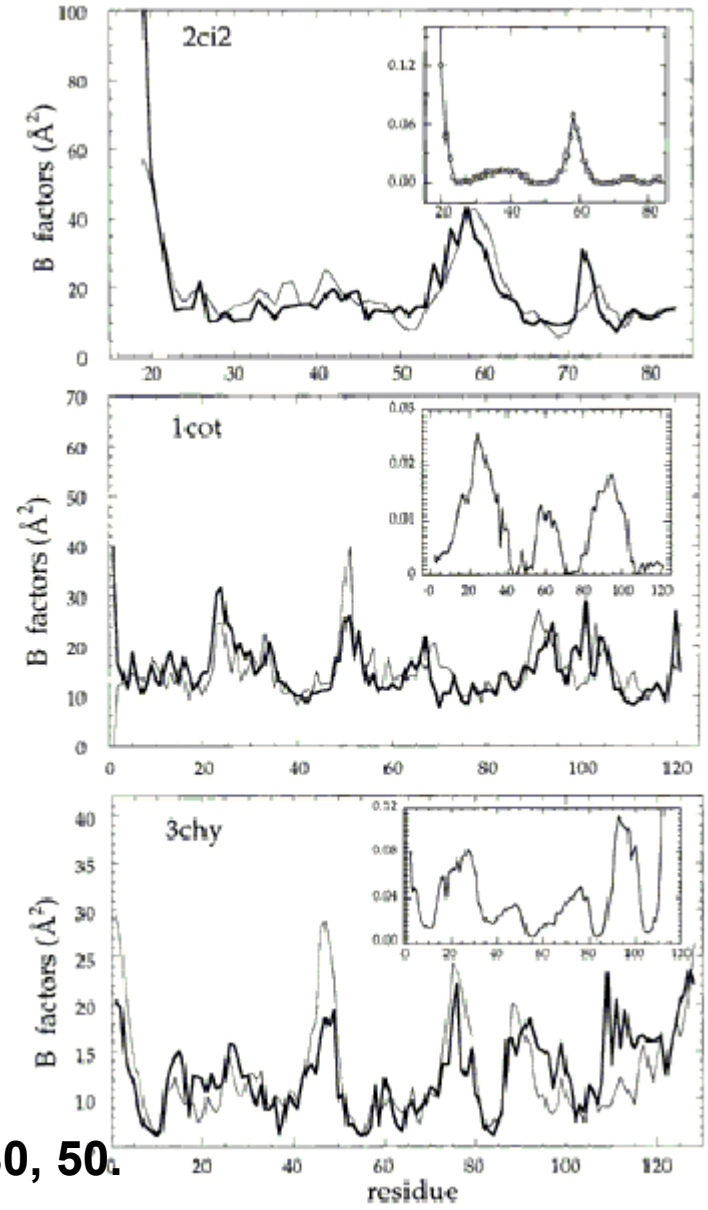
$$\Gamma_{ij} = \begin{cases} -H(r_c - r_{ij}), & i \neq j \\ -\sum_{i(\neq j)}^{N} \Gamma_{ij}, & i = j \end{cases}.$$

Demirel et al. (1998), Protein Science, v7, 2522

PENNSTATE

COMPARISON

… of theoretical (thick curve) and experimental (thin curve) B factors for Chymotrypsin inhibitor 2 (2ci2), C2 protein (1cot), and CHE-Y protein (3chy)

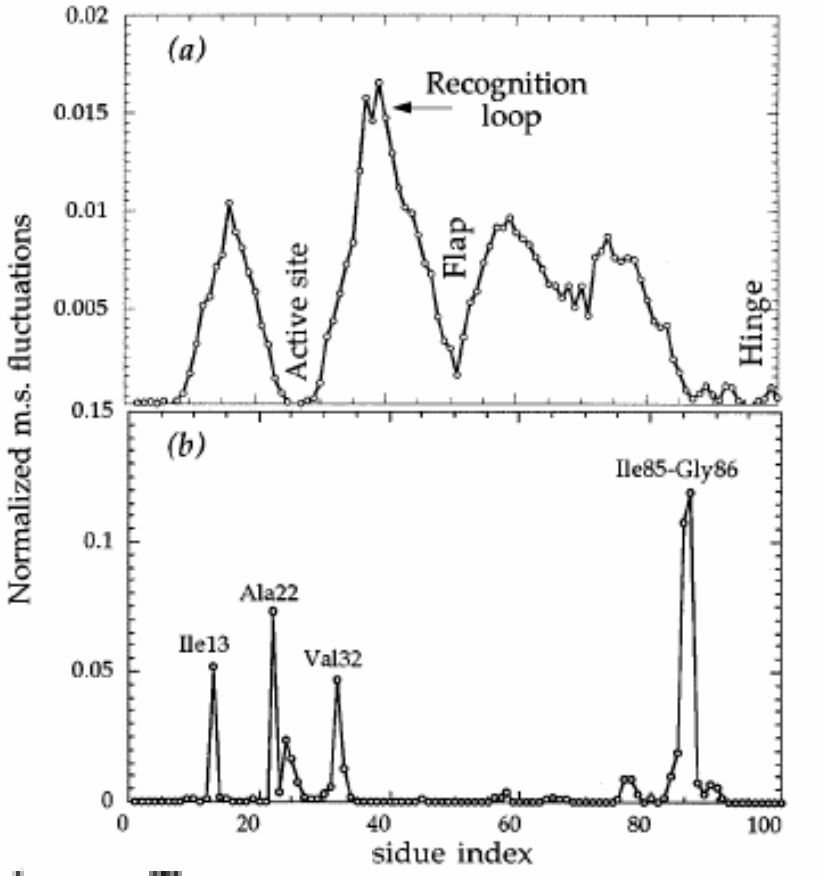


Atilgan, et al., 2001, Biophysical Journal, v80, 50.

SLOW AND FAST MOTION FUNCTION AND STABILITY

$\Gamma^{-1}=(\gamma/3k_BT)*<\Delta R\Delta R^T>$

HIV-1 protease

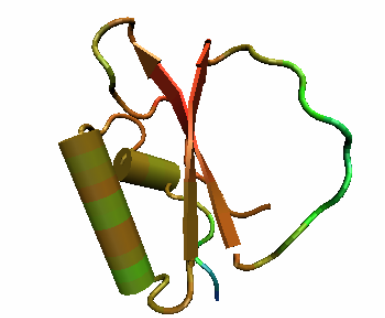


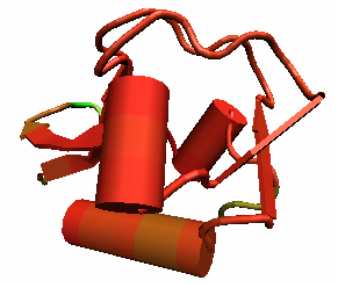
$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle_k = (3k_BT/\gamma)[\lambda_k^{-1}\mathbf{u}_k\mathbf{u}_k^T]_{ii}$$

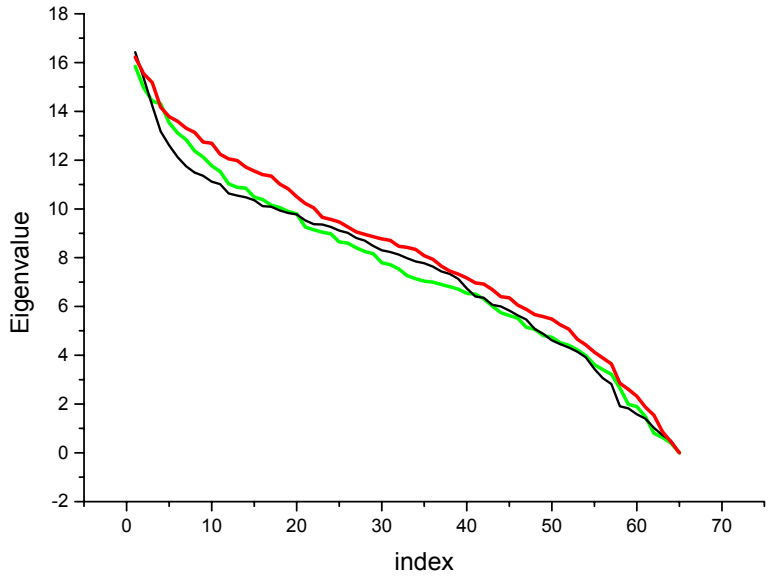
$$= (3k_BT/\gamma)\lambda_k^{-1}[\mathbf{u}_k]_i[\mathbf{u}_k]_i ,$$

Bahar, et al., (1998), Physical Review Letters, v80, 2733

Eigenvalue Distribution





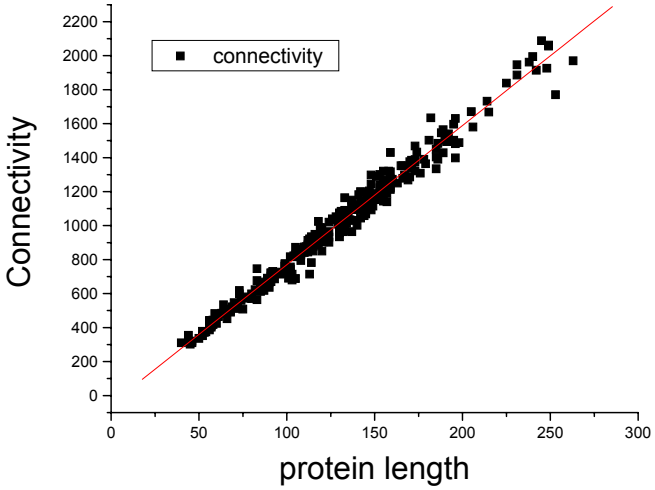


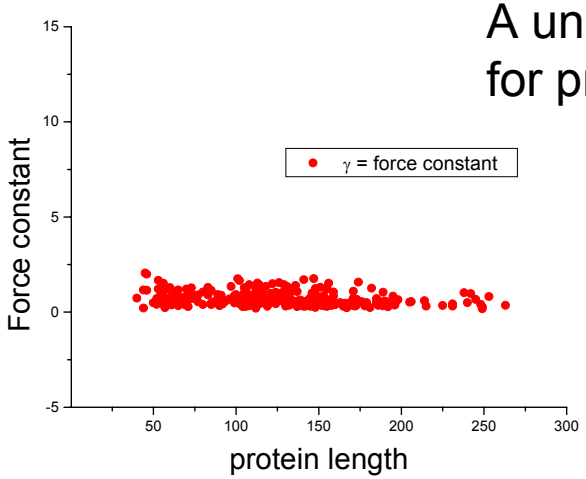
$$connectivity \equiv \sum_i \Gamma_{ii} = \sum_k \lambda_k$$

Eigenvalue distributions are similar for proteins which have equal length

Length=65

Connectivity and Force constant

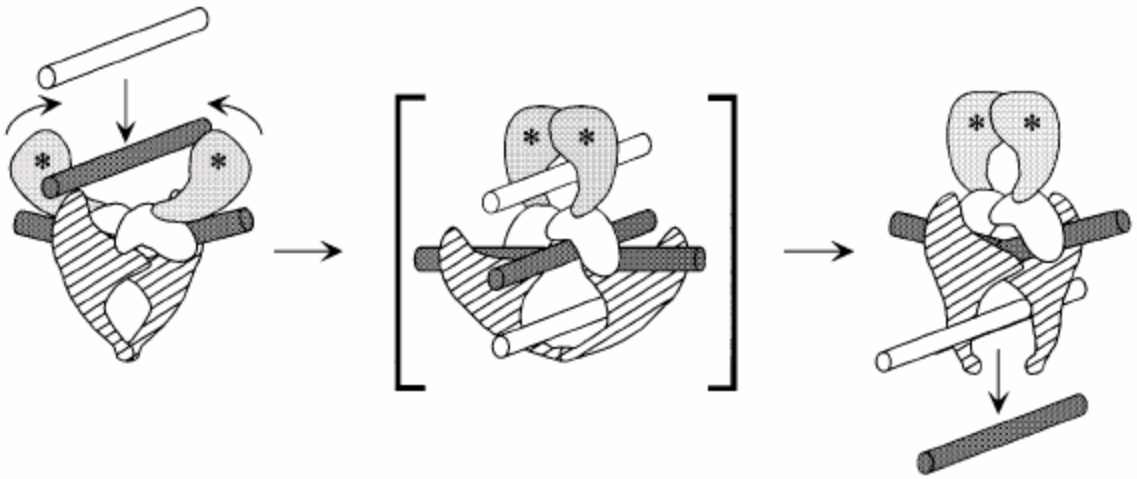




A universal constant for protein structures !

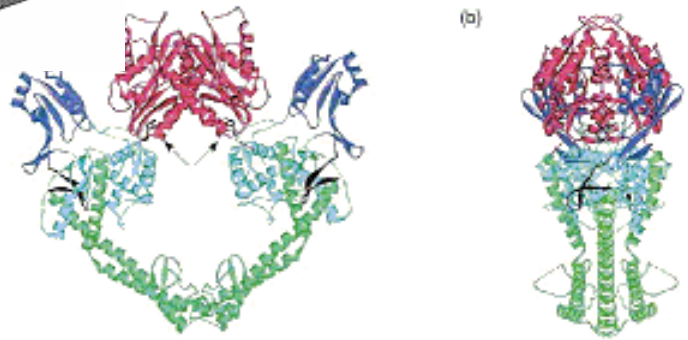
532 protein X-ray crystallography data which is extracted from PDB: connectivity is linearly varying with the protein length and γ is a constant (~1 kcal/mol Å^2)

FUNCTION

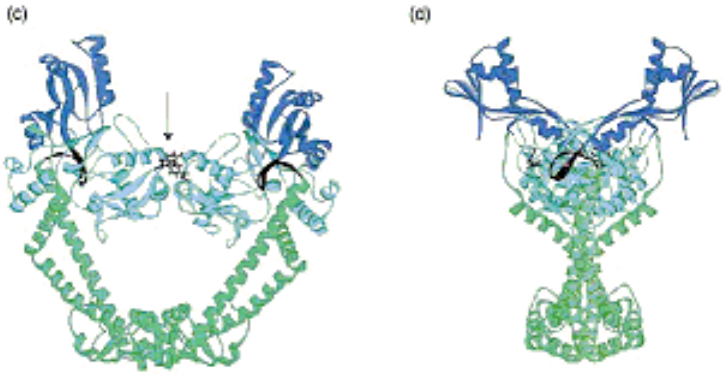


Proposed molecular mechanism for the reaction of DNA topo II.

DNA topo II and GyrA crystallographically determined fragments shown in two different views each.



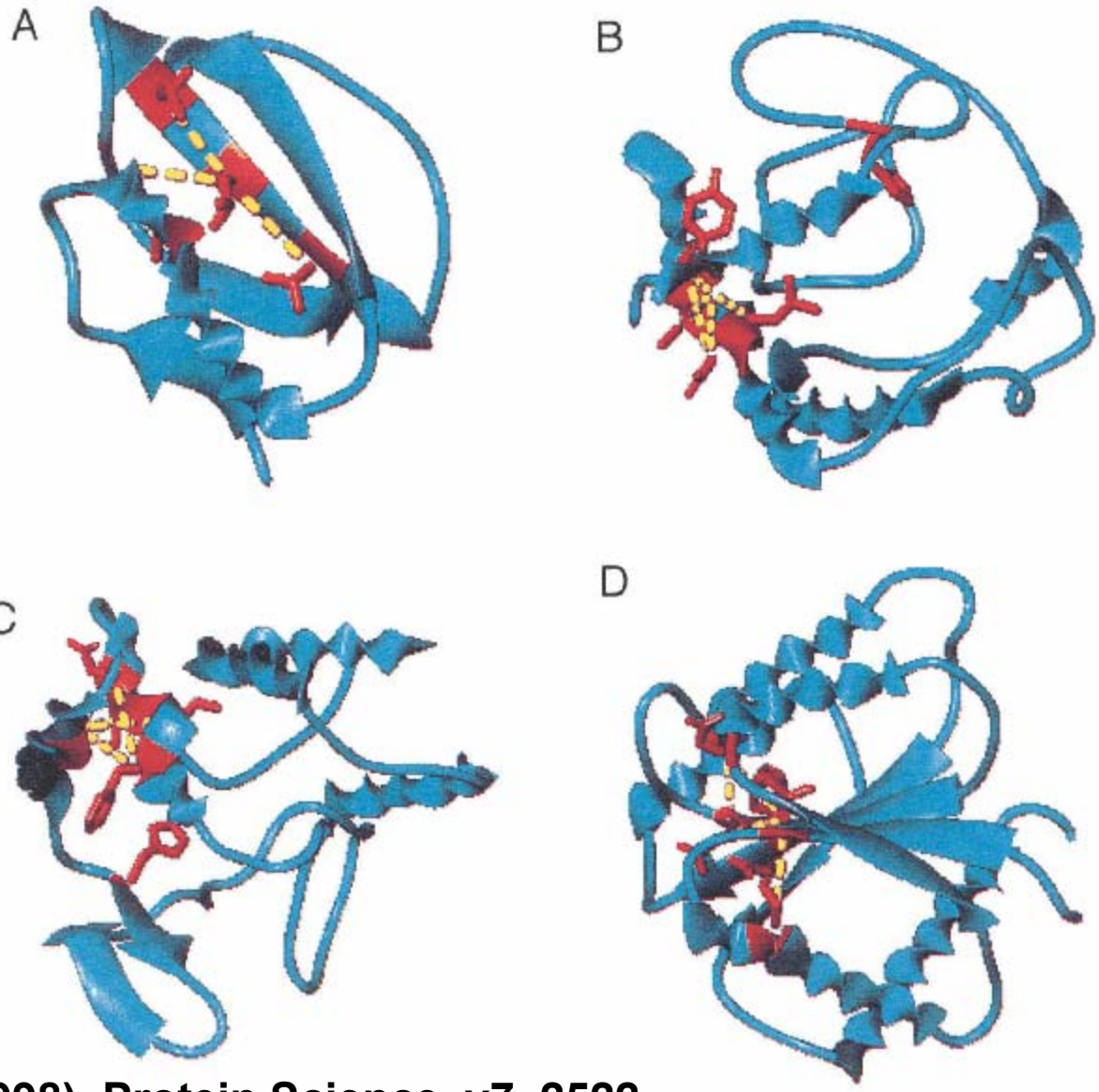
Colors denote the functional motions obtained from GNM **mode analysis**



Demirel (with others) 1999, Int. J. Quantum Chem., Vol. 75,, pp. 301

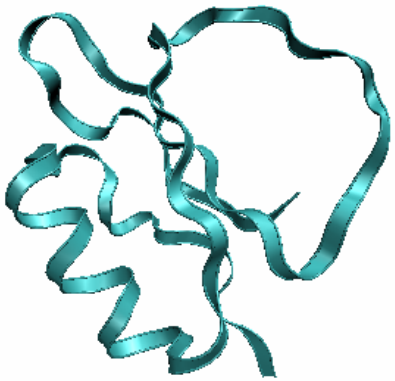
STABILITY

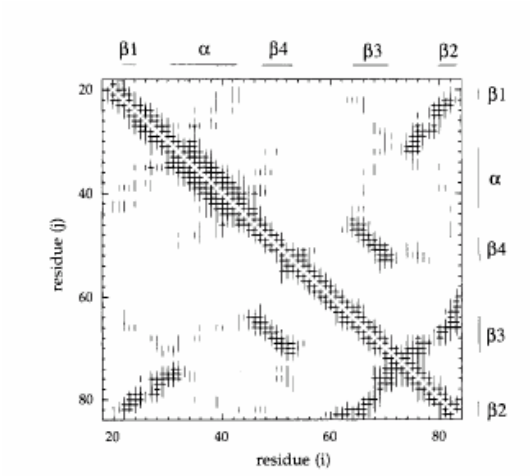
Fast modes determine stability

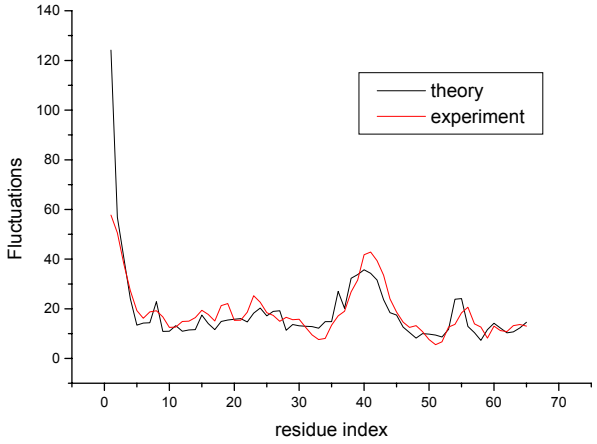


Demirel et al. (1998), Protein Science, v7, 2522

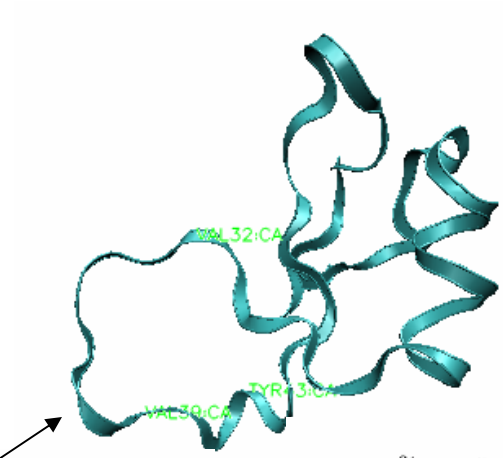
Elastic Network Model: Summary





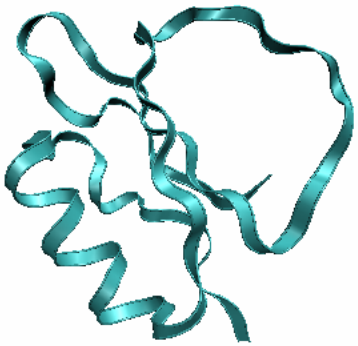


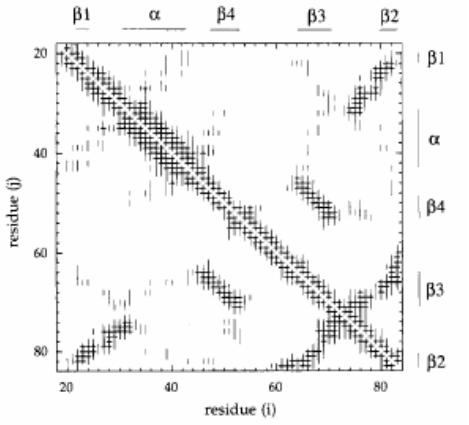
Structure → Contact Map → Fluctuation → Function, Stability

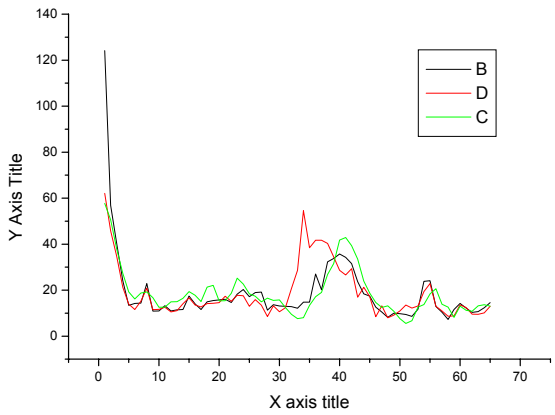


unfolded

Decoy



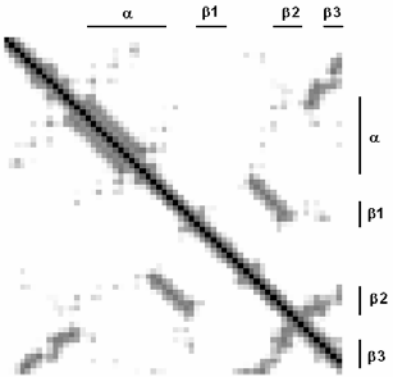


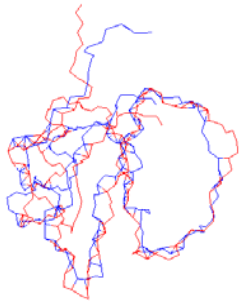


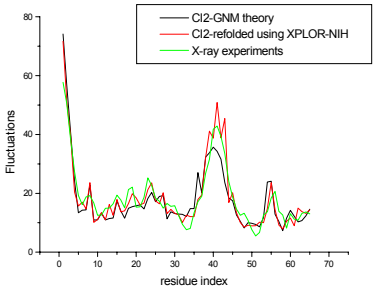
Structure ⟵ ⟶ Contact Map ⟵ ⟶ Fluctuation

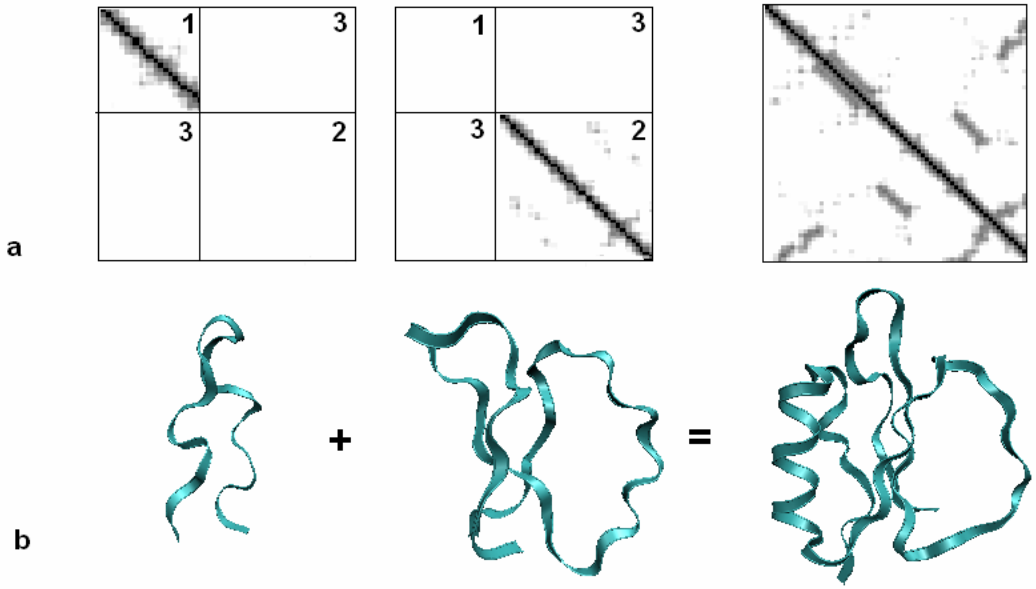
Demirel, M.C., Cherny, D., **Clustering and Diversity of Surface Fluctuations for Proteins**, submitted, 2004

“BUILDING BLOCKS” of PROTEINS

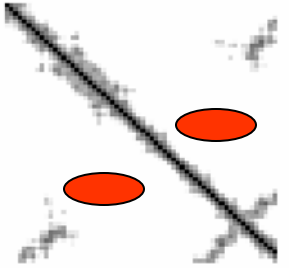


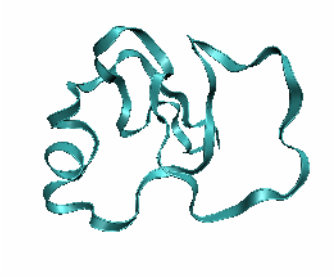


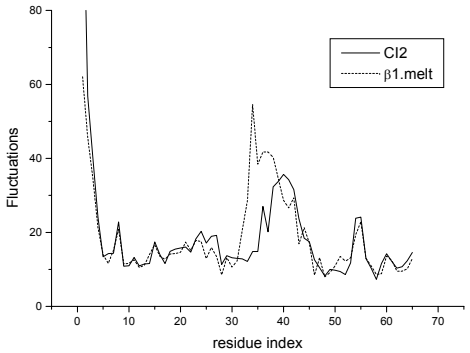


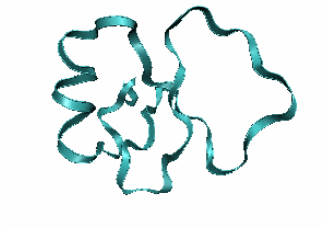


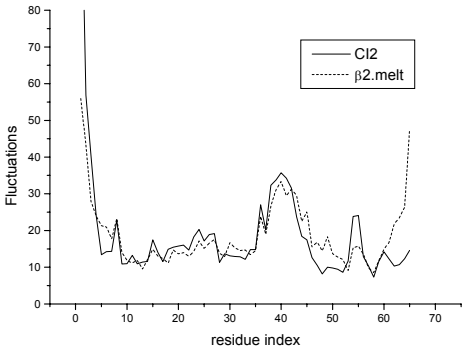
CLUSTERING OF FLUCTUATIONS

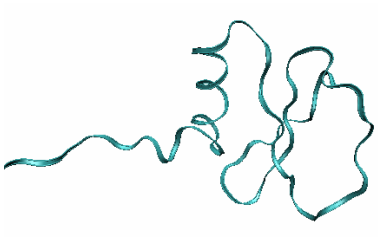


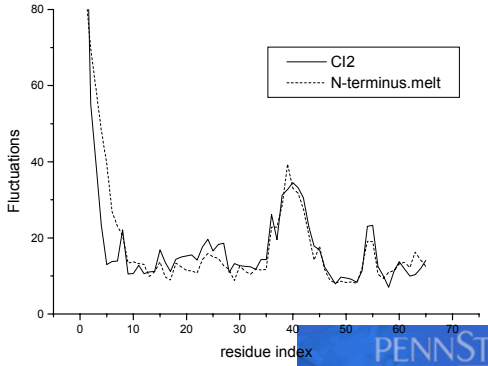




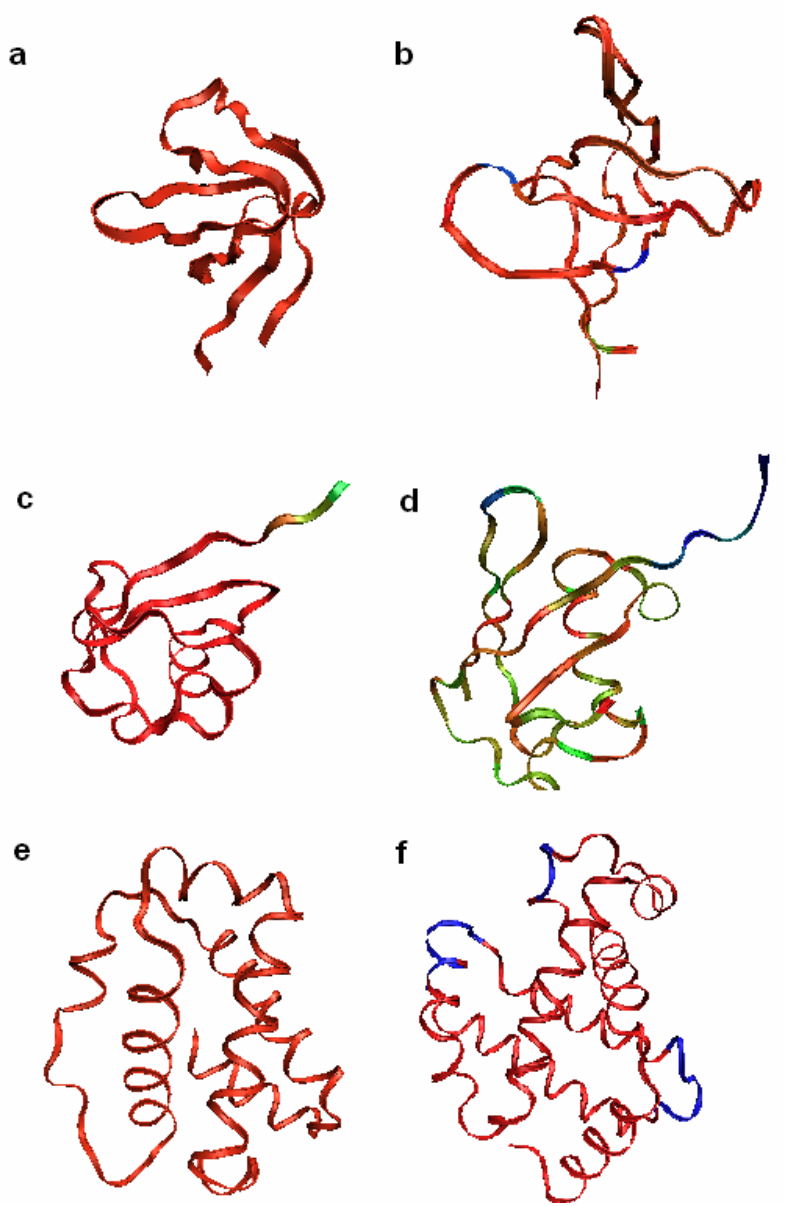








PROTEIN COMPLEXITY FROM FLUCTUATIONS



less complex more complex

The highly flexible regions are shown in blue and the least flexible with red.

SH3 domains are from nematode *C. elegans* (pdb.3sem) (a) and *H. sapiens* (pdb.5hck) (b) ; ubiquitin is from *E. coli* (pdb.1f0z) (c) and *H. sapiens* (pdb.1ubi) (d); and hemoglobin is from *Paramecium* (pdb.1dlw) (e) and *H. sapiens* subunit (pdb.1bz0, chain B) (f).

Demirel, M.C., Keskin, O., **Protein Interactions and Fluctuations in a Proteomic Network using an Elastic Network Model,** Vol 22, JBSD, 2004

PENNSTATE

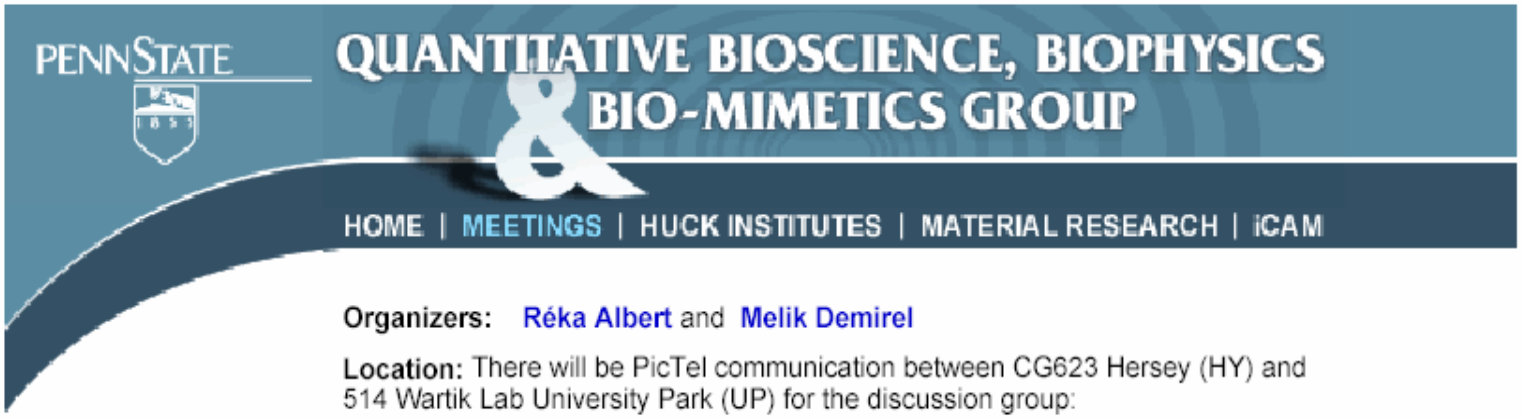
Conclusion

- Elastic network models are successful in describing equilibrium protein motions.
  - ***Experimental results (X-ray, NMR relaxation) agree well*** with this simple model.
  - Our approach introduces a new concept for classifying ***building blocks of proteins*** based on thermal fluctuations.
  - ***Large biological assemblies*** (viruses, structural proteins-titin, biomachinaries) can be simulated using elastic network models

- Protein Networks will open new avenues for biology / biophysics applications. Understanding binding of proteins will improve our conceptions about “how nature design proteins”.

PENNSTATE

Outreach: Quantitative Bioscience seminars at Penn State



Organizers: Réka Albert and Melik Demirel

Location: There will be PicTel communication between CG623 Hersey (HY) and 514 Wartik Lab University Park (UP) for the discussion group:

Time: 1:00pm-2:00pm Wednesday (see the agenda below)

How to join: If you are interested, please e-mail me (mcd18@psu.edu) so that we will add you to the list.

Click to the links on agenda for extra information and presentations (pdf)

AGENDA:

FALL'2003

19-November-2003

- Welcome and discussion: Reka Albert & Melik Demirel welcome.pdf

ACKNOWLEDGEMENT

- **GROUP**
- Murat Cetinkaya (Ph.D. Student)
- Eric So (M.S. Student)
- Lee Salway (Undergrad-EE)

- **COLLABORATORS**

-***Protein Simulations***

O. Keskin: National Cancer Institute, National Institutes of Health, MD, USA

I. Bahar: University of Pittsburgh Medical School, PA, USA

J. Sofo: Pennsylvania State University, Physics, PA, USA

-***Green Fluorescence Protein***

A. Zeytun, A. Bradbury: Los Alamos National Laboratory, NM, USA

-***Protein Assembly & Q-dots***

T. Jovin: Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

K.Kalkan: Nanofab, Pennsylvania State University, PA, USA

O. Mayans: Biozentrum, University of Basel, Basel, Switzerland

- **FUNDING**

-National Science Foundation-ICAM funding

-Penn State University, MRSEC-MRI/Huck Seed Grant

-Penn State University, Start-up funds

-Alexander von Humboldt Fellowship

Alexander von Humboldt Stiftung / Foundation

PENNSTATE